HIV and Molecular Mimicry

Gerald Myers and Hong Lu

MS K710, Los Alamos National Laboratory, Los Alamos, New Mexico 87545

For more than three decades, molecular mimicry has been considered a contributing factor in viral pathogenesis. The phenomenon comes about when viral antigens share epitopes with host cell proteins. Because tolerance is thought never to be complete, the mimicking microbe can induce anti-self antibodies, a state of autoimmunity, with resulting immunopathogenesis. This mode of attack seems preeminently possible with immunosuppressive viruses such as HSV-1 and HIV; it is also thought likely when cross-species-induced pathogenesis is involved. For some general discussions of molecular mimicry and autoimmunity brought about by viruses, see [1,2].

In the context of AIDS pathogenesis, some have argued over the years that HIV mimicry could be the crucial factor in the disease expression, for example [3–7]. If this be so, clearly the chances for a safe and successful HIV vaccine are seriously weakened [8,9]. Thus it is of considerable interest to examine mimicry claims that have been brought forward over the last decade of AIDS research. In this section, we offer a general critique of mimicry claims in so far as they have been based upon sequence similarities; experimentally-derived data will be cited on occasion, however the focus is primarily upon amino acid sequences and peptide structures. For his reason, this section appears in both the genetic and immunologic sections of the 1996 database compendium.

A. PROLEGOMENA

An important distinction that is often blurred in the literature is the distinction between *similarity* and *homology*. The correct meaning of similarity, according to the prevailing school of thought, is as an empirical relationship that can be quantified; whereas homology is an inference, a qualitative prediction about common ancestry [10–12]. Accordingly, it is impossible to prove that any two sequences are not homologous [12]. Two sequences can be similar without being homologous as a result of evolutionary convergence, and we imagine this to be the basis for viral molecular mimicry.

From this point of view, it is careless to speak of homologies in the context of mimicry, which by definition implies similarity. A pair of sequences, or a pair of structures, can be 10%, 30%, 95%... similar, but they are either homologous or they are not; philosophers speak of the first kind of judgement as determinate, whereas the judgement about homology is indeterminate. Finally, the term similarity usually makes allowances for equivalencies as well as identities: the amino acids serine and threonine may be regarded as equivalent in certain positions in a character string, and a scoring matrix would assign partial credit to one substituted for the other.

Sequences with similarity of 50% or more will typically have similar structures, although it is now well-known that two identical pentapeptides could have different shapes or presentations in space [13]. On the other hand, sequences that are marginally similar, say 30%, may have equivalent structures, thus it is not a trivial matter to show that two dissimilar sequences are also structurally dissimilar. While the emphasis herein will be upon sequence similarities, the complex relationship between sequence space and shape space must be held in mind throughout.

Moreover, it may be the case that for effective breakthrough of tolerance a mimicking epitope of a virus will not be perfectly identical to the host "self" protein. For example, an experimentally-verified mimicry between rabbit myelin basic protein (MBP) and hepatitis B virus polymerase (HBVP) is [2]:

```
MBP     TTHYGSLPQK
           YGSLPQ
HBVP    IGCYGSLPQE
```

In other cases, however, it is known that a single amino acid replacement can alter binding of an antibody to a protein [14].

We understand that some epitopes are discontinuous. Mimicry involving these epitopes will not be addressed herein. Antigen presentation is yet another complexity that will be sidestepped.

It is tempting to evaluate mimicry claims by invoking secondary structure prediction algorithms such as SOPM, discussed in an accompanying Part III section of the 1996 compendium. But we must also keep in mind long-range interactions, oligomerization, and protein residue modification (phosphorylation, glycosylation, etc.), all of which can invalidate judgements based on the primary sequence comparison and secondary structure prediction. Furthermore, we must ask whether the putative mimicking epitope on the viral antigen represents a host "self" peptide that is similarly presented on a surface [15].

It would appear from these manifold complications that mimicry can only be satisfactorily critiqued from an experimental standpoint, and indeed Oldstone emphasizes testability in his review of viral mimicry (see figure 1 in reference 2, for example). Be that as it may, untested mimicry hypotheses fill the literature of HIV—some have gained strength from time alone—and therefore critique at the sequence level might provide some insight in the absence of experimental test.

It occurs to us, in particular, that many mimicry claims made on the basis of suspected sequence similarities in the 1980's should be revisited today when the database is orders of magnitude larger. If mimicry is a pervasive factor in AIDS pathogenesis, then the sequence/structural similarity would probably be observed across many diverse variants and homologous types. In 1987, for example, a borderline similarity between HIV envelope and a molecule called neuroleukin was claimed; the story was an intriguing one in so far as it offered a basis for AIDS neuropathogenesis [16]. An experiment performed with an SIV envelope accompanied the sequence claim, but unfortunately the authors neglected to also assess the similarity between the SIV envelope and the so-called neuroleukin molecule, which turned out to be weaker than borderline. It was subsequently shown that the putative neuroleukin was, in fact, a commonplace enzyme; yet the neuroleukin hypothesis for AIDS-related neuropathy persisted in the literature after the sequences and experiment had been fully critiqued [3].

In the following section, we shall first discuss the potential contribution of compositional similarities to mimicry. Subsequent sections will include, among other things, a catalog of mimicry claims related to HIV envelope protein. At a later time, this discussion might be expanded to cover other HIV proteins, especially Gag and Nef epitopes, which are also thought to be involved in mimicry (see Part III section on Alignments, Database Searching, and Structure Prediction).

B. COMPOSITIONAL CONSIDERATIONS

One strategy for discovering similarities between HIV antigens and host proteins is to start with the amino acid composition of the former. Lentiviruses have skewed base compositions (high A, low C) and unusual codon preferences that result in significant compositional differences from what is generally seen in the PIR and Swiss-Prot protein databases [17]. HIV has a relatively high tryptophan content, for example, and tryptophan is one of the least frequently found amino acids in eukaryotes and most viruses. Polar residues are relatively high in HIV. Starting from this fact, Fitzgerald and coworkers cataloged microbial sequences that shared with HIV the unusual base composition [18]. Of these, most were bacterial virulence factors, *i.e.*, surface antigens, hence the authors speculated about similarities of immune dysfunction brought about by such diverse pathogens as protozoans, Pneumocystis, Treponema, and HIV. While the Fitzgerald study does not direct itself to HIV mimicry of host antigens (so much as it is concerned with antigen switching), the overall premise is heuristic: mimicry probably begins with gross similarities—similar charge, similar presentation, similar extent of glycosylation, etc. PROPSEARCH (http://www.heidelberg.de/aaa.html) and AACompIdent (http://expasy.chuge.ch/ch2d/aacompi.html) are two programs accessible through Internet for identifying a protein from its amino acid composition.

Taking a more focused approach from what has just been described for HIV and compositional similarities, Douvas and Sobelman conducted sweeping sequence comparisons involving three preselected sequence sets: a set of nuclear antigens, some of which are known to be involved in autoimmune disorders; a set of viral proteins, some of which are related to immunosuppression; and a set of 41 control proteins not known to have autoimmune significance [19]. Pentameric and hexameric similarities were

found to the human nuclear antigens in both the viral and control sets, but an excess of "hits" was observed for viral proteins and two nuclear antigens—the 70 kDa component of RNP particles, which is characteristically involved in mixed connective tissue disease (MCTD), and CENP-B, a centromere protein which is associated with scleroderma. Both proteins are highly hydrophilic. Several viruses, HIV-1, HSV-1, EBV, SRV-1, and CMV, were implicated. The proposed matches to HIV were to both gp120 and gp41, and because the strongest instances involved the 70 kDa splicing protein, the authors followed that up in [20].

The similarity between HIV-1 gp41 and the hydrophilic COOH-terminus of the 70 kDa ribonucleoprotein component has undeniable compositional implications, as is evident below

```
HMM Consensus (gp41)    GPDRPEGIEEEGGEQDRDR
                        GPD P G EE G ++DR R
U1-snRNP 70 kDa         GPDGPDGPEEKGRDRDRER
```

A comparable situation of sequence compositional bias is found with a stretch of the HIV/SIV Nef protein and the CENP-B centromere antigen studied in [19]:

```
HMM Consensus (Nef)     DDWDDEDEEEVGFP
                        DD D+ED +EV +P
CENP-B                  DDDDEEDGDEVPVP
```

These alignments were taken from a Smith-Waterman (S-W) search using an HMM-generated consensus for HIVs, as described in an accompanying section of Part III of the 1996 compendium. The log-odds score in each case (in bits) is about 1, which implies a two-fold greater than even probability that the similarity does not obtain from chance alone; as a rule a score should be 20 or more to have high specificity. The similarities and the marginal scores seen in these alignments stem from the low complexity (relative monotony) of the sequences involved.

The HMM Env consensus, which was generated from about 400 HIV and SIV sequences, extends the observations of [19] and [20] to primate immunodeficiency viral sequences in general, but it also reveals the ease with which such a weak match can be found: for comparison, the similarity score between the relevant region of gp41 and the *homologous* CAEV transmembrane protein is 5.5, not a high score but more than twenty-fold higher than the score implied by 1. This is not itself a judgment that the mimicry claim is false; rather, it merely points to a weakness in the sequence argument that led to, and helps sustain, the claim. Supporting the claim are the observations that: 1) the COOH-terminus of the 70 kDa can itself beget autoantibodies (which can interfere with splicing function); 2) the region of gp41 involved is known to be immunogenic; 3) similarities, however lacking in complexity, are found in clusters rather than at a single locus; 4) the authors find a 50 to 100% enhancement over control sera of binding of anti-RNP antibodies to gp41 in an ELISA assay; and 5) HIV sera show reactivity to snRNP 70 kDa in western blotting [20].

In the face of weak similarities such as those shown above, it is instructive to search a database larger than 41 control proteins to discover the ease with which matches to the 70 kDa and gp41 regions in question are encountered. Running a BLASTP search with the 70 kDa splicing protein sequence (BLAST also computes log-odds scores), more than 900 matches can be uncovered which have scores and Poisson probabilities superior to those associated with the relevant gp41 fragment. Among viruses, many represent more likely matches to 70 kDa peptides than does HIV gp41; and not all of these possibly mimicking viruses are immunosuppressive (*e.g.* human papillomaviruses). Hence, what we might call "dry mimicry" (computer-based similarity) of the 70 kDa splicing protein may be commonplace; it remains to be seen how extensive the "wet mimicry" will be.

If gp41 mimics the 70 kDa splicing protein, it might also have the potential to mimic other cellular proteins on the same basis of hydrophilicity. BLASTing the database for other matches to just the relevant C-terminal portion of the HMM consensus gp41 sequence produces an array of weak matches to several cellular proteins. One of the strongest of these borderline matches is indeed the 70

kDa RNP (Poisson probabilities between 0.07 and 0.11). This is in some contrast to the comparatively weaker score from the S-W search based on the entire HMM consensus for Env (see above).

Mimicry analyses can also be fruitfully conducted using the BLOCKS database, version 9.1, which consists of 3300 multiply aligned, gapless arrays of similarities and which features position-dependent scoring [21,22] (http://www.blocks.fhcrc.org). The snRNP family of proteins is represented in the BLOCKS database as three blocks, BL00030A,B,C, the third of which is most representative of the COOH-terminus of the 70 kDa splicing protein. One of the distinct advantages of the BLOCKS search strategy is that clusters of similarity can be identified. A search of the database using the relevant COOH-terminus of the gp41 consensus sequence did not reveal a significant match to any of the 3300 blocks. We will have reason to return to the BLOCKS search strategy below. At this point in our assessment, we have some experimental evidence in support of the claim, but the sequence arguments that underpin the claim appear to be insignificant.

Before leaving the topic of compositional bias, a brief comment about probabilities is in order. *A priori* probabilities based on the fraction 1/20 raised to some power are generally neither accurate nor useful for sequences of epitope length; they typically underestimate the likelihood of meaningless similarities, especially when compositional bias is present. In general, statistical relationships between a query sequence and target database sequences will be strongly affected by sequence length and composition, and the presumption of a normal distribution in sequence searching can distort the results [23]. Log-odds scores, used in PAM matrices, BLAST, and HMM-related database searching, are more trustworthy. For example, the HMM/S-W search strategy [24],

$$\text{score} = \log_2 \frac{P(S_i|M)}{P(S_i|R)}$$

where the alignment of each sequence in the database, S_i, is compared to both the HMM generated model, M, and a random model, R. The latter should have the same amino acid composition as the database at large and it should be as likely, *a priori*, as M. The log-odds score corrects for sequence length. With BLAST, as we have seen in a previous paragraph, Poisson probabilities are also calculated and reported alongside the log-odds score [25]. Output from a BLOCKS search includes a nomogram to assist the user in interpreting the probability of a chance similarity.

C. SEQUENCE DEGENERACY and MIMICRY

A second mimicry claim regarding HIV envelope and the 70kD splicing protein involves the well-studied gp120 V3 loop of the former and the nucleic acid binding sequence of the latter, both of which are immunogenic [20]. In this instance, low complexity, or compositional uniformity, is not obviously involved; on the other hand, widespread variability in the V3 loop calls into question the breadth of the claim:

HIV-1 IIIB	RGPGRAFVTIGK
U1-snRNP 70 kDa	KPRGYAFIEYEH
HIV-1 B subtype consensus	IGPGRAFYTTGE
	mawrkvwfar-d
	lqlkqtlhr-eq
	v qwssvv-ark
	t f ggiigidg
	f a -ryrsnkr
	k g qmwvkta
	r v xsdysnn
	s s tt yqh
	y t x pss
	m xp

The authors in [20] took note of the similarity between the HIV-1 IIIB V3 loop and the 70 kDa splicing protein, shown in the top two lines of the alignment. Shown below them is the comparison

with the B subtype consensus sequence and in lowercase letters, in descending order of frequency, the observed replacements at the various positions in B subtype loops as catalogued in Part III of the HIV database compendium. Other HIV-1 subtypes, A, D, E, O, etc., manifest variable V3 loops and HIV-2s have a very different sequence and structure altogether. This raises the question whether mimicry occurs with select variants when it doesn't occur throughout a viral type.

We do not find that the HMM consensus for the HIV/SIV V3 loop region takes part in any significant matches in a S-W/HMM search. Furthermore, a meaningful similarity is not found in a BLOCKS database, even though the 70 kDa region of interest, the nucleic acid binding sequence, is represented as block BL00030B. On the other hand, reactivities between anti-RNP sera and the IIIB V3 antigen in an ELISA assay were found to be as great as those between HIV positive sera and the IIIB loop. Furthermore, in one MCTD patient who was infected with HIV, anti-V3 and anti-70 kDa titers were observed to vary in tandem [20]. Finally, we should note that an established case of mimicry, the rabbit myelin basic protein and the hepatitis B viral polymerase discussed above, would not yield significant BLAST similarity.

We must tentatively conclude that some serological effects are being captured, yet we must also wonder how extensive they can be given the well-documented sequence degeneracy. (It is conceivable that some of the V3 loop sequences in the database were taken from defective viruses, but it is also certainly the case that the IIIB loop sequence is not the only viable sequence.) Finally, it is simply possible that the effects seen for the IIIB V3 and the 70 kDa RNP component are not biologically significant, in spite of the cross-reactivities and some sequence similarity (however weak) [26].

HIV is one of the most degenerate (variable) microbial pathogens. The previous example suggests that hypotheses based on sequences from the earliest characterized HIV-1 strains, such as IIIB, may not be indicative of a general picture of mimicry, unless viral mimicry is thought to be sometimes particular, rather than being uniformly generic. The following table of mimicry claims in the literature involving HIV envelope and host-cell proteins sequences attempts to summarize this situation. The summary is not exhaustive, but it covers the majority of claims about mimicry involving HIV gp120 and gp41 and known host protein sequences. The table lists the original claim, reports a best individual match set (first three columns), then presents the claim in light of the variation encountered in both the viral and the cell protein (fourth column). HMM-generated consensus sequences have been taken from Part II of the 1996 database compendium; their generation is discussed in a separate section of Part III. In some instances, a consensus could be deduced for the cellular protein (fourth column).

One immediately notes that a major fraction of the HIV envelope mimicry claims involve immune-related host proteins—HLA, Ig, complement, etc. This is partly an historical artifact: many of these claims were brought forward in the 1980's when an overabundance of immunoglobulin-related sequences was present in the protein libraries, hence a database search with any query sequence was likely to stumble upon borderline matches to immune-related molecules [23,27]; and in the case of an HIV sequence query, any connection to immunopathology is suggestive. This situation was of course exacerbated by the fact that the full range of HIV variation was unappreciated at the time.

Many of the claims included in the table have been critiqued at the bench by Neurath and coworkers, who simply conclude that "immunization with gp120/gp160 is unlikely to elicit harmful autoimmune responses" [8].

gp120

Source	Protein Coordinate	Sequence	Consensus	Ref
IgG1, IgG2, IgG3	88-97	NHKPSNTKVDK	nHKPSNTKVDK	28
HIV-1/gp120(DJ258A)	55-65	DAKAYDTEVHN	dakay?tevhn	
Most-likely HIV12-SIV			DAKAYDTEVHN	
HLA DR II beta	142-151	VVSTGLIHNG	vvst?(?)li?ng	29,30
Fas antigen	275-280	VQLIRN		5
HIV-1/gp120 (IBNG)	251–260	VVSTQLLLNG	vvstqlllnG	
Most-likely HIV12-SIV			VVSTQLLLNG	
HLA DR alpha	28-40	EEHVIIQAEFYLN		29
HIV-1/gp120 (2HT596.4)	270-282	EEVVIRSANFTDN	e(?)e??irsen?tnn	
Most-likely HIV12-SIV			EEIVIRSENFTDN	
HLA class I C alleles	66-69, 79-82	KYKR, RKLR		31
HIV-1/gp120 (HXB2R)	485-487,500-503	KYK, KAKR	KYK, ?(?)akr	
Most-likely HIV12-SIV			KYK, KAKR	
Fas protein	115-124	VEINCTR		6,32
HIV-1/gp120	292-301	VEINCTR	v?rnctr	
Most-likely HIV12-SIV			VEINCTR	
CD4 receptor	60-64	SLWDQ		33,6,5
HIV-1/gb120 (IBNG)	108-112	SLWDQ	slwd?q	
Most-likely HIV12-SIV			SLWDQ	
RV glycoprotein	189-199	CDIFTNSRGKR		34,35
a-cobratoxin	30-40	CDAFCSIRGKR		34,35
a-bungarotoxin	30-40	CDAFCSSRGKV		34,35
HIV-1/gp120	164-174	FNISTSIRGKV	fn?tt??rdk?	
Most-likely HIV12-SIV			FNITTEIRDKK	
TCR alpha-chain	26-46	PILILKQMCHKVRILMCISQT		6,5,36
HIV-1/gp120	212-232	PIPIHYCAPAGFAILKCNNKT	pipihyCapagfailkcndk?	
Most-likely HIV12-SIV			PIPIHYCAPAGFAILKCNDKK	
Ig light chain V region	37-48	QQHPGKAPKLVI	QQkPGkAP??iI	37
HIV-1/gp120 (NL43)	308-328	QRGPGRAFYVTI	???gpg?(??)afy??t	
Most-likely HIV12-SIV			TIGPGQAFYATG	
70 kDa splicing protein	319-330	KPRGYAFIEYEH		20
HIV-1/gp120(LAI)	316-423	RGPGRAFVTIGK	?gpg?(??)afy?(?)tg?	
Most-likely HIV12-SIV			IGPGQAFYATED	
Ig light chain V region	29-39	LLHSDGFDYLN	LL??Dg????n	37
HIV-1/gp120 (HXB2R)	453-463	LLTRDGGNSNN	lltrdgg(??????)n???	
Most-likely HIV12-SIV			LLTRDGGDNNSTN	
Ig light chain		CSTDINGYFLF		38
HIV-1/gp120 (LAI)	450-460	CSSNITGLLLT	c?snitGlllt	
Most-likely HIV12-SIV			CSSNITGLLLT	
rheumatoid antigen (70-kDa)	488-492	GGGDM		19
HIV-1/gp120 (IBNG)	462-466	GGGDM	gggdm	
Most-likely HIV12-SIV			GGGDM	

gp41

Source	Protein Coordinate	Sequence	Consensus	Ref
HLA DR II alpha	169-183	VEHWGLDQPL	vehwgld?PL	29
HIV-1/gp41 (SF1703)	588-597	VERYLKDQQL	vErylkd?qql	
Most-likely HIV12-SIV			VERYLKDQQL	
rheumatoid antigen (70-kDa)	407-413	RDRDRDR		19,20
rheumatoid antigen (70-kDa)	524-528	RDRDR		
rheumatoid antigen (70-kDa)	542-552	RDRDRDRDRDR		
HIV-1/gp41 (SF2)	739-743	RDRDR	?drdr	
rheumatoid antigen (70-kDa)	562-566	ERGRD		
HIV-1/gp41 (ELI)	737-741	ERGRD	e?drdr	
Most-likely HIV12-SIV			ERDRDR	
rheumatoid antigen (CENP-B)	428-433	EEEGGE		19
HIV-1/gp41 (SF1703)	739-744	EEEGGE	eeeGGe	
Most-likely HIV12-SIV			EEEGGE	
HLA DR II beta	19-25	NGTERVR	ngterv?	39,40,41,3
HIV-1/ gp41 (LAI)	829-835	EGTDRVI	egtDrvi	
Most-likely HIV12-SIV			EGTDRVI	
IL-2	34-39	LEHLLL		42,43
HIV-1/gp41 (LAI)	856-861	LERILL	leraLl	
Most-likely HIV12-SIV			LERALL	

D. CONCLUDING REMARKS

Molecular mimicry involving viral peptides does occur, however the extent and specificity are not easy to establish. By the nature of the problem, short, typically hydrophilic, sequences are involved, making rigorous similarity searches difficult. At a minimum, sequence analysts must critique their claims with respect to 1) compositional factors; 2) statistical significance; and 3) sequence degeneracy. Other important considerations include secondary and tertiary structure, amino acid modifications, and insignificant cross-reactivities, that is to say in vitro cross-reactivities that have no biological meaning. Whether autoimmune complications play a major or a negligible role in AIDS pathogenesis remains to be seen.

References

[1] Oldstone,M.B.A. (1987) Molecular mimicry and autoimmune disease. *Cell* **50**:819-820.

[2] Oldstone,M.B.A. (1989) Virus induced autoimmunity: molecular mimicry as a route to autoimmune disease. *J Autoimmunity* **2** (Suppl.):187–194.

[3] Bjork,R.L. Jr. (1991) HIV-1: seven facets of functional molecular mimicry. *Immunol Lett* **28**:91–95.

[4] Habeshaw,J., Hounsell,E., and Dalgleish,A. (1992) Does the HIV envelope induce a chronic graft-versus-host-like disease? *Immunol Today* **13**:207–210.

[5] Susal,C., Kropelin,M., Daniel,V., and Opelz,G. (1993) Molecular mimicry between HIV-1 and antigen receptor molecules: a clue to the pathogenesis of AIDS. *Vox Sang* **65**:10–17.

[6] Silvestris,F., Williams, R.C., Jr., and Dammacco,F. (1995) Autoreactivity in HIV-1 infection: the role of molecular mimicry. *Clin Immunol Immunopath* **75**:197–205.

[7] Bisset, L.R. (1994) Molecular mimicry in the pathogenesis of AIDS: the HIV/MHC mycoplasma triangle. *Med Hypotheses* **43**:388–396.

[8] Neurath,A.R., Strick,N., Li,Y.-Y., and Jiang,S. (1993) Improbability of harmful autoimmune responses resulting from immunization with HIV-1 envelope glycoproteins. *AIDS Res Human Retro* **9**:1195–1208.

[9] Weber, G.F. and Cantor, H. (1993) HIV glycoprotein as a superantigen. A mechanism of autoimmunity and implications for a vaccination strategy. *Med Hypotheses* **41**:247–250.

[10] Fitch, W.M. (1973) Aspects of molecular evolution. *Annu. Rev. Genetics* **7**:343–380.

[11] Pearson,W.R. (1996) Effective protein sequence comparison. In *Computer Methods for Macromolecular Sequence Analysis* (ed. R.F. Doolittle) pp. 227–258. *Methods Enzymol*. Vol. 266. Academic Press, San Diego, 1996.

[12] Doolittle, R.F. (1987) *Of URFS and ORFS: A Primer on How to Analyze Derived Amino Acid Sequences*. University Science Books, Mill Valley, California.

[13] Rost,B. and Sander,C. (1993) Prediction of protein secondary structure at better than 70% accuracy. *J. Mol. Biol.* **232**:584–599.

[14] Ivanoff,L.A., Looney, D.J., McDanal, C., Morris, J.F., Wong-Staal, F., Langlois, A.J., Petteway, S.R. Jr., and Matthews, T.J. (1991) Alteration of HIV-1 infectivity and neutralization by a single amino acid replacement in the V3 loop domain. *AIDS Res Human Retro* **7**:595–603.

[15] Quaratino,S., Thorpe,C.J., Travers,P.J., and Londei,M. (1995) Similar antigenic surfaces, rather than sequence homology, dictate T-cell epitope molecular mimicry. *Proc Natl Acad Sci USA* **92**:10398–10402.

[16] Lee,M.R., Ho,D.D., and Gurney,M.E. (1987) Functional interaction and partial homology between human immunodeficiency virus and neuroleukin. *Science* **237**:1047–1051.

[17] Myers, G, Retroviral Sequences. (1996) In *Retroviruses*, ed. by Coffin,J. Hughes,S. and Varmus,H. Appendix I. Cold Spring Harbor Laboratory Press, Cold Spring Harbor NY.

[18] Fitzgerald,D.J., Bronson,E.C., and Anderson,J.N. (1996) Compositional similarities between the human immunodeficiency virus and surface antigens of pathogens. *AIDS Res. Human Retroviruses* **12**:99–106.

[19] Douvas,A. and Sobelmen,S. (1991) Multiple overlapping homologies between two rheumatoid antigens and immunosuppressive viruses. *Proc Natl Acad Sci USA* **88**:6328–6332.

[20] Douvas,A. and Takehana,Y. (1994). Cross-reactivity between autoimmune anti-U1 snRNP antibodies and neutralizing epitopes of HIV-1 gp120/41. *AIDS Res Human Retro* **10**:253–262.

[21] Henikoff,S. and Henikoff,J.G. (1994) Protein family classification based on searching a database of blocks. *Genomics* **19**:97–107.

[22] Henikoff,J.G. and Henikoff,S. (1996) BLOCKS database and its applications. In *Computer Methods for Macromolecular Sequence Analysis*, ed. R.F. Doolittle, pp.88–105, Academic Press, San Diego.

[23] Altschul,S.F., Boguski,M.S., Gish,W., and Wooten,J.C. (1994) Issues in searching molecular sequence databases. *Nature Genetics* **6**:119–129.

[24] Eddy,S.R., Mitchison,G., and Durbin,R. (1995) Maximum discrimination hidden Markov models of sequence consensus. *J Comp. Biol*. **2**:9–23.

[25] Madden,T.L., Tatusov,R.L., and Zhang,J. (1996) Applications of network BLAST server. In *Computer Methods for Macromolecular Sequence Analysis*, ed. R.F. Doolittle, pp.131–141, Academic Press, San Diego.

[26] Nickerson, C., Luthra, H., and David, C. (1991) Antigenic mimicry and autoimmune diseases. *Int J Immunol* **7**:205–224.

[27] Claverie,J.-M. and States,D.J. (1993) Information enhancement methods for large scale sequence analysis. *Computers Chem*. **17**:191–201.

[28] Solder,B., Marschang,P., Wachter,H., Dierich,M.P., Nayyar,S., Lewin,I.V., and Stanworth,D.R. (1989). Anti-viral antibodies in HIV(HTLV-III) infection possess auto-antibody activity against a CH1 domain determinant in human IgG: possible immunological consequences. *Immunol. Lett*. **23**:9–19.

[29] Zaitseva,M.B., Moshnikov,S.A., Kozhich,A.T., Frolova,H.A., Makarova,O.D., Pavlikov,S.P., Sidorovich,I.G., and Brondz,B.B. (1992) Antibodies to MHC class II peptides are present in HIV-1 positive sera. *Scand. J. Immunol*. **35**:267–273.

[30] Young,J.A. (1988) HIV and HLA similarity. *Nature* **333**:215.

[31] Lopalco,L., De Santis, C., Meneveri,R., Longhi,R., Ginelli,E., Grassi,F., Siccardi,A.G., and Beretta,A. (1993) Human immunodeficiency virus type 1 gp120 C5 region mimics the HLA class I alpha1 peptide-binding domain. *Eur. J. Immunol*. **23**:2016–2021.

[32] Szawlowski,P.W., Hanke,T., and Randall,R.E. (1993) Sequence homology between HIV-1 gp120 and the apoptosis mediating protein Fas. *AIDS* **7**:1018.

[33] Zagury,J.-F., Cantalloube,H., et al. (1992) Striking identity between HIV-1 glycoprotein gp120 and its CD4 receptor [letter]. *Lancet* **340**:483–484.

[34] Bracci,L., Lozzi,L., Rustici,M., and Neri,P.(1992) Binding of HIV-1 gp120 to the nicotinic receptor. *FEBS Lett*. **311**:115–118.

[35] Neri,P., Bracci,L. Rustici,M. and Santucci,A. (1990) Sequence homology betweeen HIV gp120, rabies virus glycoprotein, and snake venom neurotoxins. Is the nicotinic acetylcholine receptor an HIV receptor? *Arch. Virol*. **114**:265–269.

[36] Kropelin,M., Susal,C., and Opelz,G. (1993) Cross-reactivity of anti-HIV with the CD4 receptor molecule. Proc. Int. Conf. AIDS, June 6-11, abstract PO-A20-0433.

[37] Metlas,R., Veljkovic,V., Paladini,R.D., and Ponger,S. (1991) Protein and DNA-sequence homologies between the V3 loop of human immunodeficiency virus type I envelope protein gp120 and immunoglobulin variable regions. *Biochem. Biophys. Res. Comm*. **179**:1056–1062.

[38] Kieber-Emmons,T., Jameson, B.A., and Morrow, W.J. (1989) The gp120-CD4 interface: structural, immunological and pathological considerations. *Biochem. Biophys. Acta* **989**:281–300.

[39] Golding,H., Shearer,G.M., Hillman,K., Lucas,P., Manischewitz,J., Zajac,R.A., Clerici,M., Gress,R.E., Boswell,R.N., and Golding,B. (1989) Common epitope in human immmunodeficiency virus (HIV) I-GP41 and HLA class II elicits immunosuppressive autoantibodies capable of contributing to immune dysfunction in HIV I-infected individuals. *J. Clin. Invest*. **83**:1430–1435.

[40] Golding,H., Robey,F.A., Gates, F.T., Linder, W., Beining, P.R., Hoffman, T., and Golding, B. (1988) Identification of homologous regions in human immunodeficiency virus I gp41 and human MHC class II beta 1 domain. Monoclonal antibodies against the gp41-derived peptide and patients' sera react with native HLA class II antigens, suggesting a role for autoimmunity in the pathogenesis of acquired immmunr deficiency syndrome. *J. Exp.Med*. **167**:914–923.

[41] Koshino,K., Tokano,Y., Hishikawa,T., Sekigawa,I., Takahashi,Y., and Hashimoto,H. (1995) Detection of antibodies to HIV-1 gp41 and HLA class II antigen-derived peptides in SLE patients. *Scand. J. Rheumatol*. **24**:288–292.

[42] Bost, K.L. and Pascual, D.W. (1988) Antibodies against a peptide sequence within the HIV envelope protein crossreacts with human interleukin-2. *Immunol. Invest*. **17**:577–586.

[43] Reiher,W.E., Blalock,J.E., and Brunck,T.K. (1986) Sequence homology between acquired immunodeficiency syndrome virus envelope protein and interleukin 2 [published erratum appears in *Proc Natl Acad Sci USA* **84**:564,1987]. *Proc. Natl. Acad. Sci. USA* **83**:9188–9192.